



#2

SEQUENCE LISTING

<110> David Michalovich
Matthew Alan Sims
Narjis Shaikh

<120> NOVEL COMPOUNDS

<130> GP-30088-D1

<140> 09/943,689

<141> 2001-08-31

<150> US 09/184,001

<151> 1998-11-02

<150> UK 9806221.9

<151> 1998-03-23

<150> UK 9817479.0

<151> 1998-08-11

<160> 4

<170> FastSEQ for Windows Version 3.0

<210> 1

<211> 2186

<212> DNA

<213> HOMO SAPIENS

<400> 1

cttggcacga	ggcctcgtgc	caggcggcat	gagggcccg	ggccccgggg	gctgaggcgc	60
ccgcgcctg	ccgcgggggc	cgctcgcgtc	ctccatggag	gccggagagg	aaccgctgct	120
gctggccgaa	ctcaagcccg	ggcgcccca	ccagtttgat	tggaagtcca	gctgtgaaac	180
ctggagcgtc	gccttctccc	cagatggctc	ctggtttgct	tggtctcaag	gacactgcat	240
cgtcaaaactg	atccccctggc	cgttggagga	gcagttcatc	cctaaagggt	ttgaagccaa	300
aagccgaagt	agcaaaaatg	agacgaaagg	gcggggcagc	ccaaaagaga	agacgctgga	360
ctgtggtcag	attgtctggg	ggctggcctt	cagcccgctg	ccttccccac	ccagcaggaa	420
gctctgggca	cgccaccacc	cccaagtgcc	cgatgtctct	tgcttggttc	ttgctacggg	480
actcaacgat	gggcagatca	agatctggga	ggtgcagaca	gggctcctgc	ttttgaatct	540
ttccggccac	caagatgtcg	tgagagatct	gagcttcaca	cccagtggca	gtttgatttt	600
ggtctccgcg	tcacgggata	agactcttcg	catctgggac	ctgaataaac	acggtaaaca	660
gattcaagtg	ttatcgggcc	acctgcagtg	ggtttactgc	tgttccatct	ccccagactg	720
cagcatgctg	tgctctgcag	ctggagagaa	gtcggctctt	ctatggagca	tgaggtccta	780
cacgttaatt	cggaagctag	agggccatca	aagcagtgtt	gtctcttggt	acttctcccc	840
cgactctgcc	ctgcttgtca	cggcttctta	cgataccaat	gtgattatgt	gggaccoccta	900
caccggcgaa	aggctgaggt	cactccacca	caccaggtt	gaccccgcta	tggatgacag	960
tgacgtccac	attagctcac	tgagatctgt	gtgcttctct	ccagaaggct	tgtaccttgc	1020
cacggtggca	gatgacagac	tcctcaggat	ctgggccctg	gaactgaaaa	ctcccattgc	1080
atttgctcct	atgaccaatg	ggctttgctg	cacatttttt	ccacatgggt	gagtcattgc	1140
cacagggaca	agagatggcc	acgtccagtt	ctggacagct	cctagggtcc	tgtcctcact	1200
gaagcactta	tgccggaaag	cccttcgaag	tttcctaaca	acttaccaag	tcctagcact	1260
gccaatcccc	aagaaaatga	aagagttcct	cacatacagg	actttttaag	caacaccaca	1320
tcttgtgctt	ctttgtagca	gggtaaatcg	tcctgtcaaa	gggagttgct	ggaataatgg	1380

gccaacatc	tggtcttgca	ttgaaatagc	atttctttgg	gattgtgaat	agaatgtagc	1440
aaaaccagat	tccagtgtac	tagtcatgga	tctttctctc	cctggcatgt	gaaagtcagt	1500
cttagaggaa	gagattccac	ttgcacggca	acagagcctt	acgttaaatt	ttcagtccag	1560
ttatgaacag	caagtgttga	actctttctg	cttgttttga	ttcaaagtgc	agttactgat	1620
gttgttttga	ttatgcaact	aagtaggcct	ccagagcctc	tctagtggca	gagcagctca	1680
cactccctcc	gctgggaacg	atggcttctg	cctagtacct	atccttgtgt	ttctgatgca	1740
gtggtagcat	tggttcaagt	tctctcctgc	tgtggtcaga	gttgcttcga	tggtggccaa	1800
gtgcttttct	tcttgggctc	ccttctgacc	tgcaggacag	ttttcctgga	gccatttggt	1860
atgaggtatt	aatttagctt	aactaaatta	caggggactc	agaggccgtg	ctcctgaccg	1920
atccagacac	tattactggc	tttttttttt	ttttttttaa	caatgggtgtg	catgtgcagg	1980
aaatgacaaa	tttgtatgtc	agattataca	aggatgtatt	cttaaaccgc	atgactattc	2040
agatggctac	tgagttatca	gtggccattt	attagcatca	tatttatttg	tattttctca	2100
acagatgtta	aggtacaact	gtgtttttct	cgattatcta	aaaaccatag	tacttaaatt	2160
gaaaaaaaa	aaaaaaaaa	aaaaaa				2186

<210> 2

<211> 404

<212> PRT

<213> HOMO SAPIENS

<400> 2

Met	Glu	Ala	Gly	Glu	Glu	Pro	Leu	Leu	Leu	Ala	Glu	Leu	Lys	Pro	Gly
1			5					10					15		
Arg	Pro	His	Gln	Phe	Asp	Trp	Lys	Ser	Ser	Cys	Glu	Thr	Trp	Ser	Val
			20				25					30			
Ala	Phe	Ser	Pro	Asp	Gly	Ser	Trp	Phe	Ala	Trp	Ser	Gln	Gly	His	Cys
		35				40						45			
Ile	Val	Lys	Leu	Ile	Pro	Trp	Pro	Leu	Glu	Glu	Gln	Phe	Ile	Pro	Lys
	50				55				60						
Gly	Phe	Glu	Ala	Lys	Ser	Arg	Ser	Ser	Lys	Asn	Glu	Thr	Lys	Gly	Arg
65				70				75					80		
Gly	Ser	Pro	Lys	Glu	Lys	Thr	Leu	Asp	Cys	Gly	Gln	Ile	Val	Trp	Gly
			85				90					95			
Leu	Ala	Phe	Ser	Pro	Trp	Pro	Ser	Pro	Ser	Arg	Lys	Leu	Trp	Ala	
	100					105					110				
Arg	His	His	Pro	Gln	Val	Pro	Asp	Val	Ser	Cys	Leu	Val	Leu	Ala	Thr
	115				120						125				
Gly	Leu	Asn	Asp	Gly	Gln	Ile	Lys	Ile	Trp	Glu	Val	Gln	Thr	Gly	Leu
	130				135					140					
Leu	Leu	Leu	Asn	Leu	Ser	Gly	His	Gln	Asp	Val	Val	Arg	Asp	Leu	Ser
145			150					155					160		
Phe	Thr	Pro	Ser	Gly	Ser	Leu	Ile	Leu	Val	Ser	Ala	Ser	Arg	Asp	Lys
			165					170				175			
Thr	Leu	Arg	Ile	Trp	Asp	Leu	Asn	Lys	His	Gly	Lys	Gln	Ile	Gln	Val
	180					185						190			
Leu	Ser	Gly	His	Leu	Gln	Trp	Val	Tyr	Cys	Cys	Ser	Ile	Ser	Pro	Asp
	195					200					205				
Cys	Ser	Met	Leu	Cys	Ser	Ala	Ala	Gly	Glu	Lys	Ser	Val	Phe	Leu	Trp
	210				215						220				
Ser	Met	Arg	Ser	Tyr	Thr	Leu	Ile	Arg	Lys	Leu	Glu	Gly	His	Gln	Ser
225				230					235					240	
Ser	Val	Val	Ser	Cys	Asp	Phe	Ser	Pro	Asp	Ser	Ala	Leu	Leu	Val	Thr
			245					250					255		
Ala	Ser	Tyr	Asp	Thr	Asn	Val	Ile	Met	Trp	Asp	Pro	Tyr	Thr	Gly	Glu
	260					265					270				
Arg	Leu	Arg	Ser	Leu	His	His	Thr	Gln	Val	Asp	Pro	Ala	Met	Asp	Asp
	275					280					285				

Ser Asp Val His Ile Ser Ser Leu Arg Ser Val Cys Phe Ser Pro Glu
 290 295 300
 Gly Leu Tyr Leu Ala Thr Val Ala Asp Asp Arg Leu Leu Arg Ile Trp
 305 310 315 320
 Ala Leu Glu Leu Lys Thr Pro Ile Ala Phe Ala Pro Met Thr Asn Gly
 325 330 335
 Leu Cys Cys Thr Phe Phe Pro His Gly Gly Val Ile Ala Thr Gly Thr
 340 345 350
 Arg Asp Gly His Val Gln Phe Trp Thr Ala Pro Arg Val Leu Ser Ser
 355 360 365
 Leu Lys His Leu Cys Arg Lys Ala Leu Arg Ser Phe Leu Thr Thr Tyr
 370 375 380
 Gln Val Leu Ala Leu Pro Ile Pro Lys Lys Met Lys Glu Phe Leu Thr
 385 390 395 400
 Tyr Arg Thr Phe

<210> 3
 <211> 2558
 <212> DNA
 <213> HOMO SAPIENS
 <220>
 <221> UNSURE
 <222> (146) (161) (178) (233) (254) (296)

<400> 3
 gaaaggaacc gctgctgctg gccgaactca agccccggcg cccccaccag tttgattgga 60
 agtccagctg tgaaacctgg agcgtcgctt tctccccaga tggctcctgg tttgcttggt 120
 ctcaaggaca ctgcatcgct aaactnatcc cctggccggt ngaggagcag ttcattccnta 180
 aagggtttga agccaaaagc cgaagtagca aaaatgagac gaaagggcgg ggnagtccaa 240
 aagagaagac gctngactgt ggtcagattg tctgggggct ggccttcagc ctgtgncttt 300
 cccacccag caggaagctc tgggcacgcc accaccccca agtgcccgat gtctcttgcc 360
 tggttcttgc tacgggactc cacgatgggc agatcaagat ctgggaggtg cagacagggc 420
 tctgctttt gaattcttcc ggccaccaag atgtcgtgag agatctgagc ttcacaccca 480
 gtggcagttt gatttttggt tccgcgtcac gggataagac tcttcgcac tgggacctga 540
 ataaacacgg taaacagatt caagtgttat cgggccacct gcagtgggtt tactgctggt 600
 ccattctccc agactgcagc atgctgtgct ctgcagctgg agagaagtcg gtctttctat 660
 ggagcatgag gtccctacag ttaattcgga agctagaggg ccattcaaagc agtggtgtct 720
 ctgtgtgactt ctcccccgac tctgccctgc ttgtcacggc ttcttacgat accaatgtga 780
 ttatgtggga cccctacacc ggcgaaaggc tgaggctcact ccaccacacc cagggtgacc 840
 ccgccatgga tgacagtga gtccacatta gctcactgag atctgtgtgc ttctctccag 900
 aaggcttgta ccttgccacg gtggcagatg acagactcct caggatctgg gccctggaac 960
 tgaaaactcc cattgcattt gctcctatga ccaatgggct ttgctgcaca ttttttccac 1020
 atgggtggagt cattgccaca gggacaagag atggccacgt ccagttcttg acagctccta 1080
 gggtcctgtc ctactgaag cacttatgcc ggaaagccct tcgaagtttc ctaacaactt 1140
 accaagtcct agcactgcca atccccaaaga aaatgaaaga gtctctcaca tacaggactt 1200
 tttaagcaac accacatctt gtgcttcttt gtagcagggt aaatcgctct gtcaaaggga 1260
 gttgctggaa taatgggcca aacatctggt cttgcattga aatagcattt ctttgggatt 1320
 gtgaatagaa tgtagcaaaa ccagattcca gtgtactagt catggatctt tctctccctg 1380
 gcatgtgaaa gtcagtctta gaggaagaga ttccacttgc acggcaacag agccttacgt 1440
 taaattttca gtccagttat gaacagcaag tgttgaactc tttctgcttg ttttgattca 1500
 aagtgcagtt actgatgttg ttttgattat gcaactaagt aggcctccag agcctctcta 1560
 gtggcagagc agctcacact cctccgctg ggaacgatgg cttctgccta gtacctatcc 1620
 ttgtgtttct gatgcagtgg tagcattggt tcaagttctc tctgctgtg gtcagagttg 1680
 cttcgatgtt ggccaagtgc ttttcttctt gggctccctt ctgacctgca ggacagtttt 1740
 cctggagcca tttggtatga ggtattaatt tagcttaact aaattacagg ggactcagag 1800
 gccgtgctcc tgaccgatcc agacactatt actggctttt tttttttttt ttttaacaatg 1860

gtgtgcatgt	gcaggaaatg	acaaatttgt	atgtcagatt	atacaaggat	gtattcttaa	1920
accgcatgac	tattcagatg	gctactgagt	tatcagtggc	catttattag	catcatattt	1980
atttgatatt	tctcaacaga	tgtaaggta	caactgtgtt	tttctcgatt	atctaaaaac	2040
catagtactt	aaattgaaca	gttgcaaaga	tgtcttaatt	gtgtaaagaa	ttggtgtagt	2100
catgacttta	gctgatactc	ttatgtacga	gatctgtctc	tgctgtttta	cttcattgga	2160
ttaatcagct	ggtttcaact	ctactgcgaa	acaaaaatag	ctccttaaaa	gtactgttct	2220
ccttcagtgg	catgtagtta	tctaatacag	acacctcatt	caaacaaaac	ctgccttagg	2280
aaaatttaat	atattttaaa	ttattttaaa	agaaatacaa	catcttattc	tttagctttc	2340
ttaatcggtg	ctttatggag	gccagtgtaa	cgttacatga	ctcgttgaga	aagttgagga	2400
atttcctcta	ccacctttgt	tgcttgaaga	aaaacatgtc	ttttcaaaat	gagaggcttt	2460
cattgaagaa	aagaaaaaaa	caacagttaa	aagctattgg	ctctctgttt	catttttttc	2520
cattaagaaa	aaaaaaagtc	ccctttttaa	acaagcag			2558

<210> 4
 <211> 400
 <212> PRT
 <213> HOMO SAPIENS
 <220>
 <221> UNSURE
 <222> (53) (59) (98)

<400> 4

Lys	Glu	Pro	Leu	Leu	Leu	Ala	Glu	Leu	Lys	Pro	Gly	Arg	Pro	His	Gln
1			5						10					15	
Phe	Asp	Trp	Lys	Ser	Ser	Cys	Glu	Thr	Trp	Ser	Val	Ala	Phe	Ser	Pro
			20					25					30		
Asp	Gly	Ser	Trp	Phe	Ala	Trp	Ser	Gln	Gly	His	Cys	Ile	Val	Lys	Leu
		35					40				45				
Ile	Pro	Trp	Pro	Xaa	Glu	Glu	Gln	Phe	Ile	Xaa	Lys	Gly	Phe	Glu	Ala
	50					55					60				
Lys	Ser	Arg	Ser	Ser	Lys	Asn	Glu	Thr	Lys	Gly	Arg	Gly	Ser	Pro	Lys
65					70					75				80	
Glu	Lys	Thr	Leu	Asp	Cys	Gly	Gln	Ile	Val	Trp	Gly	Leu	Ala	Phe	Ser
				85					90					95	
Leu	Xaa	Leu	Ser	Pro	Pro	Ser	Arg	Lys	Leu	Trp	Ala	Arg	His	His	Pro
			100					105					110		
Gln	Val	Pro	Asp	Val	Ser	Cys	Leu	Val	Leu	Ala	Thr	Gly	Leu	His	Asp
		115					120					125			
Gly	Gln	Ile	Lys	Ile	Trp	Glu	Val	Gln	Thr	Gly	Leu	Leu	Leu	Leu	Asn
	130					135					140				
Leu	Ser	Gly	His	Gln	Asp	Val	Val	Arg	Asp	Leu	Ser	Phe	Thr	Pro	Ser
145					150					155				160	
Gly	Ser	Leu	Ile	Leu	Val	Ser	Ala	Ser	Arg	Asp	Lys	Thr	Leu	Arg	Ile
				165					170					175	
Trp	Asp	Leu	Asn	Lys	His	Gly	Lys	Gln	Ile	Gln	Val	Leu	Ser	Gly	His
			180				185					190			
Leu	Gln	Trp	Val	Tyr	Cys	Cys	Ser	Ile	Ser	Pro	Asp	Cys	Ser	Met	Leu
	195						200					205			
Cys	Ser	Ala	Ala	Gly	Glu	Lys	Ser	Val	Phe	Leu	Trp	Ser	Met	Arg	Ser
	210					215					220				
Tyr	Thr	Leu	Ile	Arg	Lys	Leu	Glu	Gly	His	Gln	Ser	Ser	Val	Val	Ser
225					230					235				240	
Cys	Asp	Phe	Ser	Pro	Asp	Ser	Ala	Leu	Leu	Val	Thr	Ala	Ser	Tyr	Asp
				245					250					255	
Thr	Asn	Val	Ile	Met	Trp	Asp	Pro	Tyr	Thr	Gly	Glu	Arg	Leu	Arg	Ser
			260				265						270		
Leu	His	His	Thr	Gln	Val	Asp	Pro	Ala	Met	Asp	Asp	Ser	Asp	Val	His

	275		280		285										
Ile	Ser	Ser	Leu	Arg	Ser	Val	Cys	Phe	Ser	Pro	Glu	Gly	Leu	Tyr	Leu
	290					295					300				
Ala	Thr	Val	Ala	Asp	Asp	Arg	Leu	Leu	Arg	Ile	Trp	Ala	Leu	Glu	Leu
305					310					315					320
Lys	Thr	Pro	Ile	Ala	Phe	Ala	Pro	Met	Thr	Asn	Gly	Leu	Cys	Cys	Thr
				325					330						335
Phe	Phe	Pro	His	Gly	Gly	Val	Ile	Ala	Thr	Gly	Thr	Arg	Asp	Gly	His
			340					345					350		
Val	Gln	Phe	Trp	Thr	Ala	Pro	Arg	Val	Leu	Ser	Ser	Leu	Lys	His	Leu
		355					360					365			
Cys	Arg	Lys	Ala	Leu	Arg	Ser	Phe	Leu	Thr	Thr	Tyr	Gln	Val	Leu	Ala
	370					375					380				
Leu	Pro	Ile	Pro	Lys	Lys	Met	Lys	Glu	Phe	Leu	Thr	Tyr	Arg	Thr	Phe
385					390					395					400